

## Data-Driven Career Placement Examination System with Prediction Model in Forecasting Licensure Performance Using Regression Techniques

Louell Cid Cabanela

Polytechnic University of the Philippines, Graduate School / Sta. Mesa Manila 1016

Philippine State College of Aeronautics, AvSSM Department / Piccio Garden, Villamor Pasay City 1309

### ABSTRACT

Education plays a vital role in the development of a country, and predicting the students' performance is essential to identify future risks they might encounter and enable academic institutions to take corrective actions to prevent them from failure. This study used the descriptive and developmental method of research, and criterion sampling was used to identify/select the individuals who can provide the best information for the objective of this study. After gathering the Career Placement Exam (CPE) results, the output is now imported to the developed predictive data analysis tool on which the simple-linear regression is used. Since the CPE results are not strong enough to verify the predicted result, all the undergraduate semestral grades are also used and subdivided for each of the seven technical subjects/areas, where a multilinear regression model is used. Overall, regarding Security, Functionality, Usability, Reliability, and Portability, the level of acceptance for the developed prototype system is Highly Acceptable. Moreover, for the result of the level of accuracy using the simple linear regression model (for the CPE) and the multilinear regression model (for the seven technical areas), the accuracy level of  $\geq 85$  is based on the predicted and actual data generated in the Analytics tool. Using the equation/model derived from linear regression techniques, the machine learning prototype can determine whether the students can pass or fail the CAAP Licensure Examination as follows: if  $\alpha \leq 79.99$ , then the student will fail; if  $85.00 \leq \alpha \leq 80.00$ , then it is questionable for the student to pass and if  $\alpha \geq 85.01$  then the student will likely pass the licensure examination.

**Keywords:** Career Placement Examination (CPE); Machine learning; Multilinear regression; Predictive model; Simple linear regression


### 1. INTRODUCTION

Measuring the students' academic achievements across different subjects or fields makes a significant impact on educators to study more about different learning styles to improve not only their mode of teaching but also to lessen the factors affecting the students' performance and attain higher marks during assessments. Fortunately, the advancement of research and development benefited not only the industry sectors but also academic institutions, computer science and information technologies, manifesting the importance of predictive models (Kalechofsky, 2016).

\*Corresponding Author: [cabanela001@gmail.com](mailto:cabanela001@gmail.com) / <https://orcid.org/0000-0001-7590-9567>

Received: 02 October 2023; Received in the Revised form: 24 October 2023; Accepted: 02 November 2023;

Available online: 07 November 2023

This work is licensed under 

©2023 All rights reserved United Frontiers Publisher



Meanwhile, education plays a vital role in the development of a country, and predicting students' performance is essential to identify future risks they might encounter (Hamsa et al., 2016). Many related studies have focused on explaining and predicting learners' performance (Sorour et al., 2015). Accurate predictions of students' academic performance at the early stages of the degree program help identify weak students and enable management to take corrective actions to prevent them from failure (Pandey et al., 2016).

Some of the reasons why students' performance is essential claimed criteria for a high-quality university based on its excellent record of academic achievements in today's very competitive environment and not to waste both money and resources due to failure (Baars et al., 2017; Mesarić & Šebalj, 2016; Shahiri et al., 2015). Providing manual examinations is more prone to errors and biases, making computer-aided examination a perfect choice that can release examination results in record time and without error gaining an assessment modality Jawaid et al., (2014); Abass et al., (2017); Broughton (2017) with this some sighted as well that it cut costs than the manual or traditional way based on the study of Rønning (2017) this computer-aided examination refers to the use of computers to assess students' progress (Chalmers & McAusland, 2002).

This study aimed to develop a Data-Driven Career Placement Examination System with Prediction for Licensure Examination Performance Using a Regression Algorithm. This study highlights the developed portable mock reviewer or career placement examination system that will serve as an alternative way or mode of preparing academic institutions and their students to gain a high probability of obtaining more passers, especially hence providing sufficient guidance and time allotment to review and prepare more if necessary and also to make way for those who like to fail an opportunity for appropriate actions and interventions that will help him pass the actual exam. This study focused on the developed on-premises system that can run portably via a local server and, after generating results, will be imported on the Predictive Analytics tool forming the developed data-driven mock review/career placement exam to forecast the score/percentage using regression techniques since the numerical predictive variable has been gathered from the examination itself. This will also be raw data loaded on the predictive model analytics tool. Specifically, the study sought to answer the following questions:

1. What are the processes and challenges encountered when conducting the institution's career placement examination?
2. What is the level of acceptance in the developed Data-Driven Career Placement Examination System with Prediction for Licensure Examination Performance Using Regression Algorithm in terms of a. Security; b. Functionality; c. Usability; d. Reliability; and e. Portability?
3. What is the level of accuracy of the developed Regression Models in Forecasting Licensure Performance? And lastly,
4. What appropriate features can be designed for the system that will automate the mock examination to forecast licensure performance?

## 2. LITERATURE REVIEW

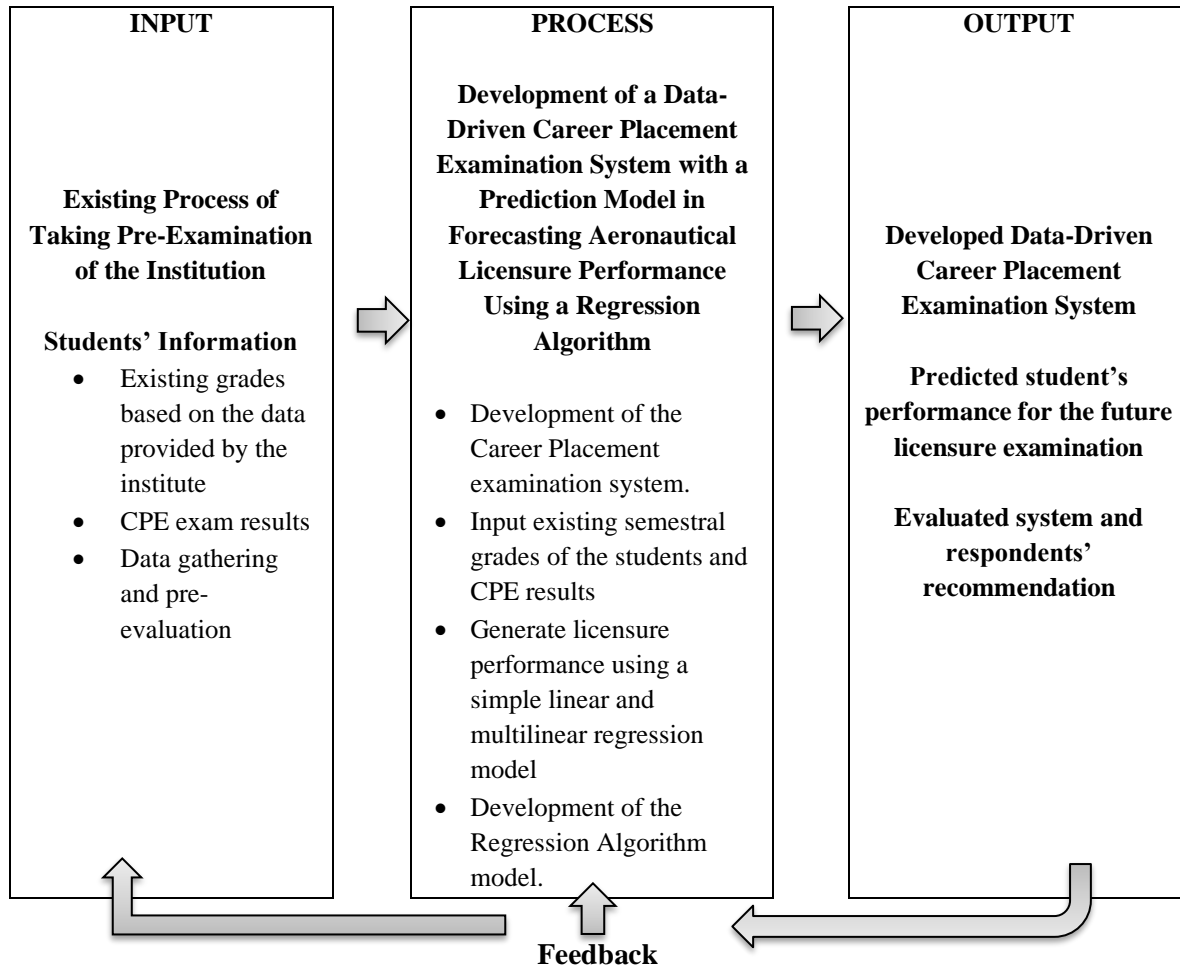
Historical data on the licensure performance of the students helps academic institutions to maintain and track their overall performance; in light of this, there are also some factors affecting their performance

as a whole for appropriate actions needed to make them more prepared and improve the number of licensure passers. Moreover, board examinations are given by professional regulatory agencies in various countries to its citizens to ensure that the desired efficiency of the job is achieved, especially those which are critical to the society like the Civil Aviation Authority of the Philippines (CAAP) that issues Aviation Maintenance Technician license for AMT graduates of Philippine State College of Aeronautics (PhilSCA) and other Aviation Institutions in the Philippines nationwide. The assessment also aims to measure the person's competencies and abilities in performing their jobs in the industry. Also, forecasting board examination results to set a prognosis of the institution's efficiency in delivering instruction nowadays can be done using a data mining technique (Abaya et al., 2016). With the advancement of technology, Computer-based exams (CBE) have several significant advantages compared to traditional paper-based exams (PBE), such as efficiency, immediate scoring and feedback in the case of multiple-choice question exams (Samson, 2017). The use of the said technology opens the opportunity for immediate feedback and support on the pre-board exam performance of the reviewer and eventually remedial before taking the actual licensure examination. This would help them cope with success and help the institution improve its licensure examination performance rating. This is another essential way of strengthening the institution's accomplishments and may eventually lead to the path toward transformation to cope with the developing information and communications technology culture (Tarun, 2017).

## **2.1 Utilization and Choosing Predictive Variables**

Choosing what to use and how to use a predictive variable is critical in data modelling because some studies are using mixed method approaches like student performance in e-learning and blended learning approach (Rakic et al., 2020) (Lu et al., 2018). Another study, like Angeles (2020), aimed to identify variables that best predict the teacher education graduates in Cagayan Philippines, namely the instruction and curriculum, to assess the quality of teaching, which only means that a predictor is a very relevant factor in forecasting performances. As a result of the prediction, not only for students but for teachers also. On the other hand, other related studies describe the different predictive variables they used for predicting student performance. In the study conducted by Son and Fujita (2019), they introduced MIMO SAPP or Multi-Output student academic performance prediction to predict the future performance of the students. They used representative sets with neural-fuzzy logic. Mock examination results as raw data for predicting the actual licensure performance are impressive. According to the study of Tarun (2017), another study also developed this kind of prototype with a DSS decision support system (Jain, 2016) (Tarun et al., 2014). However, limited studies were conducted on the development and "Integration" of Mock review examinations and the prediction of licensure performance in one system that will forecast how likely the student will pass or fail before the actual licensure exam.

## 2.2 Conceptual Framework



**Figure 1: Conceptual Framework**

The conceptual framework used in the study is the Input Process Output (IPO) model. The IPO model provides the general structure and guide for the direction of the study. By identifying and analyzing the variables of this study on the IPO model, this study came up with the conceptual framework, which is presented in Figure 1.

Figure 1 presents the conceptual framework of the study. In the input stage, this study conducted a survey questionnaire to get the demographic profile and the raw data needed, such as semestral grades/academic performance of their technical subjects and the CPE examination results. The process stage includes developing the career placement examination system, loading gathered student information, and testing the developed system results using regression techniques. For the result of this IPO model, the system could be deployed responsible for conducting mock examinations and predicting the student's actual licensure performance based on the input and process provided. The feedback represents the capability of the system to return to previous stages when errors occur before the system produces such information.

### 2.3 System Architecture

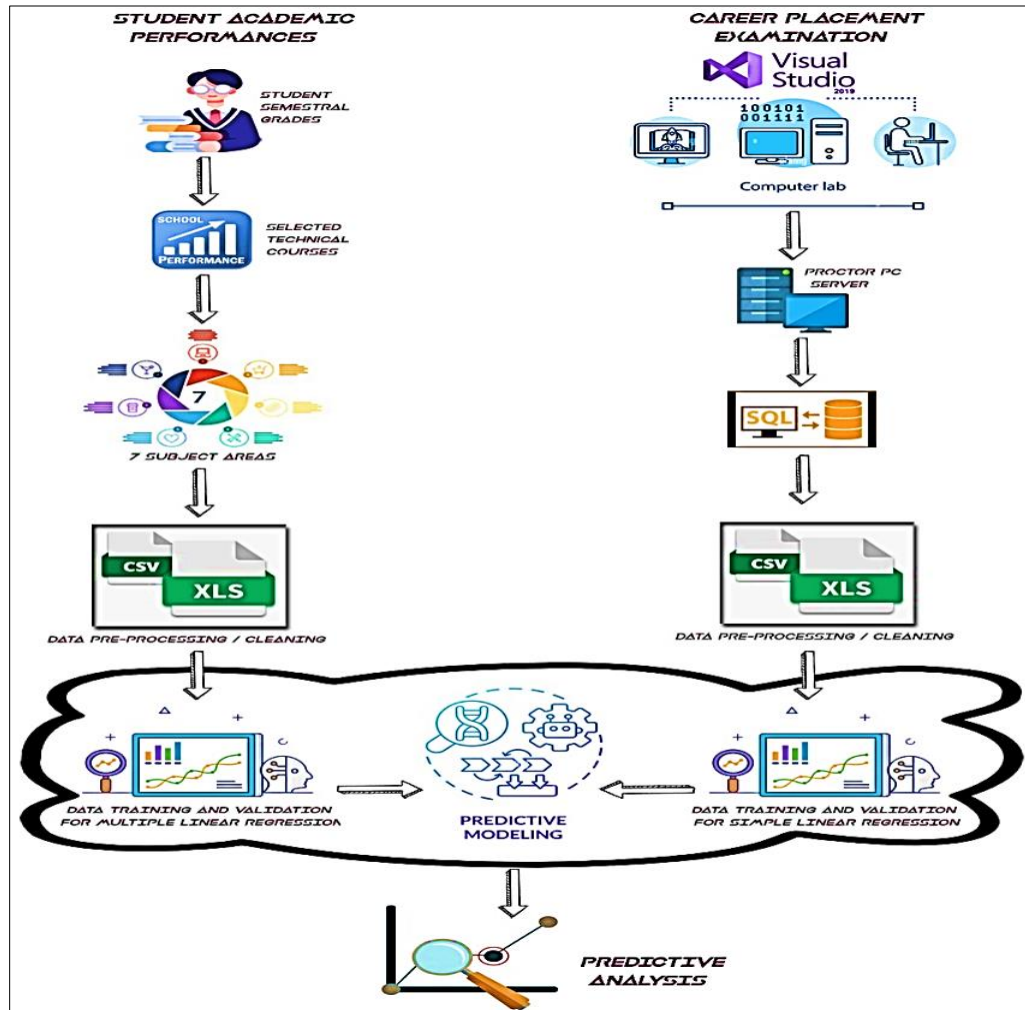


Figure 2: System Architecture of the Career Placement Examination System with Prediction Model

Figure 2 shows the interaction of the user and the Data-Driven Career Placement Examination System for Institute of Engineering students that will serve as their mock review exam for the actual licensure examination. It shows how the system works between the administrator and the examinee through MS SQL database as a standard and most widely used open-source language that is especially suited for relational database development. The datasets or the descriptive data were extracted from the 7 Technical Subject Areas and the Career Placement Exam by procedure. The raw data underwent data cleaning for pre-processing of the data and removing outliers to prevent the error output when doing the ETL (Extract Transform and Loading). Regarding the Testing Training and Data Validation, this study used a 70:30 ratio (70 percent training set and 30 percent for validation and testing dataset) because this study only used a small amount of dataset, which was quantified into 33 descriptive data. Data modeling output and analytics part will be generated from Analytics Cloud for the accuracy of the predictive data model. For the Predictive Analytics, this study used 10-fold cross-validation ( $k=10$ ), in which the data was equally divided into ten parts (9 parts

for training and one-tenth for testing) to provide a good balance between the variance and bias on the data (Zacharski, 2015).

## 2.4 Simple Linear Equation

Using the linear regression, the equation to predict if the students can pass the Licensure Examination in CAP is expressed as:

$$y = \alpha + \beta_1 x \quad / \quad y = 0.6x + 33$$

### Where:

- $y$  is the response (prediction)
- $.06x$  is the constant value (derived from the given linear equation)
- 33 is the y-intercept value

## 2.5 Multilinear Equation

Mathematically, the multilinear relationship is reflected as follows:

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n$$

### Where:

- $y$  is the response
- $\beta$  values are called the model coefficients. These values are “learned” during the model fitting/training step.
- $\beta_0$  is the intercept
- $\beta_1$  is the coefficient for  $X_1$  (the first feature)
- $\beta_n$  is the coefficient for  $X_n$  (the nth feature)

## 3. METHODOLOGY

### 3.1 Research Design

This study used descriptive and developmental research methods to accomplish this study.

### 3.2 Research Instrument / Data Collection

The research instrument used was based on the standard ISO 9126 for software development (Umar et al., 2019). Data gathering through a survey questionnaire was done first to know the current status of providing mock examination and the respondent's CPE results used to craft the system. After gathering the CPE results, the output was imported to the developed model, where the simple-linear regression was used. Since the CPE results were not strong enough to verify the predicted result, this study also included the past semestral grades for the seven areas (namely: Air laws and Airworthiness, natural Science and Aircraft general knowledge, Aircraft engineering, Aircraft maintenance Human Performance, Airframe and Powerplant), these areas stated are the composition of the mock review examination of the institute is utilized for the multilinear regression algorithm.

### 3.3 Sampling and Sample Size

For the descriptive part, which is the first part, this study used criterion sampling for the study respondents who answered the challenges and the level of acceptance as they were responsible for conducting

the career placement examination at the institution. There are total of 5 (five) pool of CPE end users in PhilSCA composed of 1 MIS-Head or IT Expert, three departmental proctors that are program coordinators from the Engineering department (1 for the Aircraft Maintenance Technology department, 1 for the Aviation Electronics Technology department and 1 for Aeronautical Engineering department) also one representative from the head of student affairs/ guidance services unit who annually facilitated the Mock review examination since they are the primary beneficiary of the developed system. The second part, which is the dataset, is composed of 33 AMT alums/graduates that contains the undergraduate historical data such as the career placement examination results from the institution and the academic performances (all semestral grades for all the semesters they took) where this study did the pre-processing part to distribute and organize all of the technical subjects for each of the seven areas according to CAAP ratings.

### 3.4 Ethical Statement

This study ensured the confidentiality of the respondents, especially the 33 AMT alums/graduates' datasets who were selected as part of data modeling and who gave their academic records voluntarily via Google Forms with informed consent that the data collected would be used for research purposes only in compliance to the Philippine Data Privacy Act of 2012 which states on Republic Act 10173 to protect the fundamental human right of privacy, communication while ensuring free flow of information. Regarding the dignity of the respondents, this study prioritizes not to disclose the names of the individuals/groups who participated in the study.

### 3.5 Analysis Techniques

This study used mean (average value) to measure the central tendency for the challenges faced by the institution and also the ISO 9126 for the developed prototype. The data modeled were trained and validated through machine tools such as Python and Weka. Lastly, to test the level of accuracy, the CPE results were tested first to see if it is correlated to the CAAP RATING and then loaded into the SAP Analytics cloud to verify/check their accuracy.

## 4. RESULTS AND ANALYSIS

To describe the respondents and support the sampling method used, it was divided into 2 (two parts) The first is the descriptive part, which consists of 5 pools of CPE users (100 percent total of administrators conducting the mock examination in the institution). The second part is the dataset consisting of 33 AMT alums/graduates. Table 1 below shows the distribution of the respondents of the study for the pool of CPE users:

**Table 1. Percentage Distribution of Respondents**

<b>Respondents</b>	<b>Frequency</b>	<b>Percentage</b>
Proctor for 3 Departments	3	60%
MIS personnel	1	20%
Guidance personnel	1	20%
<b>Total</b>	<b>5</b>	<b>100%</b>

Table 1 shows the percentage distribution of the pool of CPE users who annually conducted the career placement examination for Aircraft Maintenance Technology students under the Institute of Engineering and Technology.

#### 4.1. Existing Practice and Challenges Encountered in Conducting the Career Placement Examination in The Institution

**Table 2. Existing Practice and encountered challenges in the institution**

Encountered Challenges	Frequency
There is no automated examination for the mock examination.	5
Checking the mock examination is very time-consuming.	5
The results for the mock examination take approximately 3 to 5 days.	5
The proctor does not focus on the examinee's mistakes or subjects that need more attention (for review).	5
Currently, no employed program in the institution will track their weaknesses so that students can be more prepared for the incoming licensure students' performance and focus on their weaknesses to be more prepared for the licensure examination.	5

Table 2 shows the Existing Practice and encountered challenges of the institution. The results show that the existing practice of the institution is the PBE (paper-based examination), which includes a manual or traditional way of conducting, checking and evaluating the examination and based on their responses, there is no automated system or software program used for the CPE of the institution which resulted to the delay of publishing / posting the announcement for the mock exam passers which undeniably will take time when generating reports if the proctor or the person in charge will check them manually or individually. Moreover, the institution cannot monitor the examinee's weaknesses in the subjects or areas the students should have focused on for the incoming actual licensure examination.

#### 4.2. Acceptability Level in the Developed Data-Driven Career Placement Examination System with Prediction for Licensure Examination Performance Using Regression Algorithm

**Table 3. Level of Acceptability of the Developed System**

	N	Mean	Interpretation
Security	5	3.90	<b>Highly Acceptable</b>
Functionality	5	3.80	<b>Highly Acceptable</b>
Usability	5	3.64	<b>Highly Acceptable</b>
Reliability	5	3.30	<b>Moderately Acceptable</b>
Portability	5	3.76	<b>Highly Acceptable</b>

Table 3 shows that in terms of Security, 3.90 General Weighted mean, which is verbally interpreted as Highly Acceptable, where the proposed system has an adequate security feature that helps the user and the administrator to prevent any possible unauthorized access by having the unique system-generated examination ID password. In terms of Functionality, with a 3.80 General Weighted Mean, which is verbally interpreted as Highly Acceptable based on the findings, the proposed CPE system can provide the actual system-generated report/performance right after taking the mock examination, distribution of questions from the database with the randomized questionnaire distribution feature is also added. In terms of Usability, 3.64 General Weighted Mean, which is verbally interpreted as Highly Acceptable that the proposed system can generate real-time results, which is needed for posting the examinee passers; however, based on the result,



they are often confused about how to fully use the system that the examinees might need the help and supervision of the proctor during the actual examination. In terms of Reliability, 3.30 General Weighted Mean, which is verbally interpreted as Moderately Acceptable, the result shows that the evaluation of MIS-IT Expert and the Pool of CPE End Users is moderately acceptable in general for the reason that during the pilot testing of the proposed system, there are some technical problems using different operating systems. In terms of Portability, 3.76 General Weighted Mean, which is verbally interpreted as Highly Acceptable, the pool of users liked the plug-and-play version of the proposed system and the most appreciated feature because it can be easily installed on a Windows operating system and can also be used as a personal reviewer.

#### 4.3. Level of Accuracy of the Developed Regression Models in Forecasting Licensure Performance.

**Table 4. Test of Accuracy results using Simple Linear Regression Technique**

	Ave CPE	CAAP Ratings	Prediction	Accuracy
<i>ds1</i>	88	90	85.80	97.50
<i>ds2</i>	84	84	83.40	99.29
<i>ds3</i>	85	85	84.00	98.82
<i>ds4</i>	86	86	84.60	98.37
<i>ds5</i>	88	85	85.80	97.50
<i>ds6</i>	83	80	82.80	99.76
<i>ds7</i>	88	92	85.80	97.50
<i>ds8</i>	86	89	84.60	98.37
<i>ds9</i>	88	86	85.80	97.50
<i>ds10</i>	85	85	84.00	98.82
<i>ds11</i>	85	84	84.00	98.82
<i>ds12</i>	86	84	84.60	98.37
<i>ds13</i>	87	84	85.20	97.93
<i>ds14</i>	77	85	79.20	102.86
<i>ds15</i>	83	84	82.80	99.76
<i>ds16</i>	85	83	84.00	98.82
<i>ds17</i>	83	82	82.80	99.76
<i>ds18</i>	83	80	82.80	99.76
<i>ds19</i>	76	85	78.60	103.42
<i>ds20</i>	84	86	83.40	99.29
<i>ds21</i>	80	84	81.00	101.25
<i>ds22</i>	85	88	84.00	98.82
<i>ds23</i>	87	86	85.20	97.93
<i>ds24</i>	83	86	82.80	99.76
<i>ds25</i>	87	87	85.20	97.93
<i>ds26</i>	80	85	81.00	101.25
<i>ds27</i>	84	90	83.40	99.29
<i>ds28</i>	83	84	82.80	99.76
<i>ds29</i>	93	85	88.80	95.48
<i>ds30</i>	87	84	85.20	97.93
<i>ds31</i>	78	85	79.80	102.31
<i>ds32</i>	83	84	82.80	99.76
<i>ds33</i>	90	91	87.00	96.67

*Legend:	
<i>ds</i>	is the dataset number
Ave CPE	is the Average Result of the CPE
CAAP Ratings	is the target actual predicted value

Table 4 shows the target results of the third column, the actual CAAP ratings. Using the simple linear regression, the developed equation to predict if the students can pass the Licensure Examination in CAAAP is  $y=0.6x+33$ . This equation predicts the CAAP Ratings (Prediction =  $0.6 \times \text{AveCPE} + 33$ ). A Pearson r moment correlation was conducted to test the significant relation of the Career Placement Examination Scores of the participants to the Passing Rate they obtained in the CAAP License.

For the simple linear regression model to determine the accuracy of the prediction: Accuracy = Prediction / Average CPE \* 100 (this formula should be used for individual prediction). Based on the output, there are several factors why the developed model cannot be perfected. Some of it is due to psychological factors, examinees' condition during the examination proper, lack of preparedness, and mental conditional state while taking the mock review examination. These are only some of the reasons why the prediction overlaps. In addition, the Ave CPE result is the basis for predicting the CAAP ratings (target). As a result, the Ave CPE values are close to the Prediction values rather than the CAAP Ratings based on the derived output/outcome.

**Table 5. Test of Accuracy results using Multiple Linear Regression Technique**

	ALA	NSAGK	AE	AM	HP	AF	PP	Prediction	Accuracy
<i>ds1</i>	90.43	85.21	89.15	89.55	88.75	88.25	91.17	86.77	97.57
<i>ds2</i>	84.57	88.57	81.85	77.45	86.00	81.75	82.50	82.63	99.27
<i>ds3</i>	74.29	90.86	86.00	86.73	87.00	85.13	85.83	85.10	99.98
<i>ds4</i>	89.29	86.50	86.31	90.91	87.75	87.75	88.83	87.68	99.42
<i>ds5</i>	88.14	83.79	87.08	89.00	87.50	87.25	89.50	86.17	98.52
<i>ds6</i>	87.57	83.71	85.38	84.36	89.25	84.25	86.17	83.51	97.31
<i>ds7</i>	91.86	83.64	88.85	90.45	89.50	89.13	92.33	86.72	97.01
<i>ds8</i>	87.71	85.29	91.38	90.36	89.00	88.88	90.67	86.73	97.40
<i>ds9</i>	88.29	84.86	87.77	87.91	87.75	89.38	86.33	86.30	98.67
<i>ds10</i>	86.57	80.14	82.77	85.45	86.25	83.50	85.00	83.74	99.40
<i>ds11</i>	82.14	82.64	82.62	83.45	88.00	82.50	81.83	82.36	98.86
<i>ds12</i>	80.71	87.29	82.23	84.73	87.50	90.13	85.67	84.74	99.15
<i>ds13</i>	87.29	87.71	84.23	85.82	89.00	85.75	86.33	85.23	98.43
<i>ds14</i>	85.57	89.50	83.46	86.18	86.00	86.13	90.00	86.53	99.82
<i>ds15</i>	84.00	82.36	81.31	83.18	81.75	80.13	79.83	84.31	103.07
<i>ds16</i>	83.14	84.86	82.38	81.00	85.00	80.13	87.14	82.83	99.34
<i>ds17</i>	87.00	88.14	87.92	88.64	91.25	88.75	90.17	85.89	96.68
<i>ds18</i>	86.43	87.36	86.15	85.82	88.50	85.38	84.50	85.21	98.73
<i>ds19</i>	91.00	81.43	89.23	91.09	91.00	89.63	92.00	85.83	96.07
<i>ds20</i>	82.57	88.21	83.77	84.45	81.50	81.25	85.17	86.21	102.82
<i>ds21</i>	83.43	85.07	85.00	83.73	84.50	85.25	86.83	84.71	99.86
<i>ds22</i>	89.29	89.79	88.85	90.73	90.25	92.13	91.00	88.12	97.59
<i>ds23</i>	85.29	85.50	87.77	87.18	80.75	86.88	87.33	87.83	102.34
<i>ds24</i>	80.29	84.36	80.31	82.36	78.50	81.00	81.17	85.13	104.92
<i>ds25</i>	86.57	84.64	90.69	88.64	91.25	92.13	89.67	85.39	95.85
<i>ds26</i>	85.14	87.21	85.08	86.73	87.25	81.63	86.67	85.25	99.51
<i>ds27</i>	88.43	86.07	90.15	89.27	90.25	89.63	88.33	86.32	97.12
<i>ds28</i>	82.00	89.07	85.00	83.64	80.25	83.38	85.83	86.73	103.04
<i>ds29</i>	85.86	85.86	83.77	87.36	89.25	84.25	85.33	84.86	98.72
<i>ds30</i>	84.00	86.36	84.69	83.45	84.25	84.75	92.00	84.95	99.19
<i>ds31</i>	85.43	84.00	87.46	86.09	91.00	89.25	87.17	83.89	96.20
<i>ds32</i>	88.00	84.00	83.85	85.82	86.25	85.00	84.67	85.25	99.86
<i>ds33</i>	91.00	90.00	87.31	89.91	88.50	88.63	88.17	88.34	99.18

*Legend:	
<i>ds</i>	is the dataset number
ALA	is the Air Laws & Airworthiness
NSAGK	is the Natural Science & Aircraft General Knowledge
AE	is the Aircraft Engineering
AM	is the Aircraft Maintenance
HP	is the Human Performance
AF	is the Airframe
PP	is the Powerplant
Prediction	is the predicted values
Accuracy	individual accuracy of the prediction

Table 5 shows the prediction results based on the dataset for each of the 7 Technical Subject Areas through their accumulated grades. A significant regression was found ( $F(7,27)=2.769$ ,  $p<.026$ ). The participants' predicted passing rate through their accumulated grades is equal to this equation:

$$\text{Prediction} = 39.57 + (.146 * \text{Air Laws \& Airworthiness}) + (.254 * \text{Natural Science \& Aircraft General Knowledge}) + (.021 * \text{Aircraft Engineering}) + (.358 * \text{Aircraft Maintenance}) + (-.337 * \text{Human Performance}) + (.113 * \text{Airframe}) + (-.018 * \text{Powerplant})$$

For the multilinear regression model to determine the accuracy of the prediction:  $\text{Accuracy} = \text{Prediction} / \text{Average Rating for 7 Areas} * 100$  (this formula should be used for individual prediction). For the overall Accuracy:  $\text{Percent Accuracy} = \text{Accurate Prediction} / 33 * 100$  (this formula should be used for the overall CPE Accuracy level), the inaccurate ratings are those who exceed the 100 per cent Accuracy Level individually. In light of this, some factors prevent the model from attaining a higher accuracy. It is because of inconsistencies in the semestral grades when analyzed manually from individual raw data/ academic performances. One of the reasons is the instructor's judgement when giving the final grades for the students, not to mention the different academic styles and strategies when conducting assessments for academic performances for the factors that underlie laboratory periods, quizzes, significant exams, seat works, assignments, group related activities, research/project-based assessments and other teaching/learning pedagogical strategies that may contribute on the overall accuracy of the developed multiple linear regression model.

#### Likelihood of passing the actual examination:

**Table 6. The alpha values and their interpretation utilizing the Prediction Results**

Alpha Value	Interpretation
$\alpha \leq 79.99$	Failed
$85.00 \leq \alpha \leq 80.00$	Questionable
$\alpha \geq 85.01$	Passed

The derived alpha values in Table 6 above may help aspiring graduates who want to take the licensure examination for Aircraft Maintenance Technician. The alpha values above were based on the passing rate standard for CAAP licensing ratings. The main objective is to consider the examinee's current status/progress after taking the career placement examination system. This can help him/her identify which aspect of the examination he/she needs to focus more on to gain a high probability of passing the AMT licensure examination.

#### 4.4. Features designed for the career placement examination system that will automate the mock examination used to forecast licensure performance

##### 4.4.1. Admin Module

In this software module, the Main Menu Interface is the default screen for the administrator and proctor. The following are the features inside this module: The manage tab feature consists of Users (both administrators and the examinee's accounts). This allows the administrator to add, edit, and save new administrators/proctors to use the system. In this feature, the administrator/proctor can also add, edit and manage examinee accounts. Examinee Records is the feature that contains all of the examinees who are uploaded on the system before, during, and after the examination (here, the user/ admin can add, edit, save

and view the result before and after taking the examination for each examinee). The View Result feature will show the overall exam results, including the item/s they missed or failed to answer, which leads to the realization/decision to take corrective action/s on what to do if the examinee fails (needs to retake/ review and study more to pass the examination). The time limit feature allows the administrator/proctor to manage the time allotment for each question and adjust the time allotted depending on the weight and difficulty of the question. The randomized Item feature will apply to the system after selecting the questions from the Questionnaires to the Question bank. This feature randomized all the questions from the question bank distributed via the local server during the examination. The questionnaire loading feature allows the proctor/admin to load, encode, add, edit, and delete a specific question to the system. The question bank feature consists of all the questions that will be loaded/taken by the examinees, where the proctor/administrator selects from the bulk load of the questionnaire feature. The view tab feature consists of reports to be printed by the system. It consists of all records uploaded on the system, including the print stub before taking the exam, those examinees who passed and those who failed. Report View feature shows all examinee records, such as those who already took/did not take the exam, and identifies how many passed and failed the examination on their remarks. The Print Stub feature shows the examinee details, such as the Examinee ID and Password, before taking the examination and suggested to be distributed individually for each examinee to avoid improper use of the examinee account. The Print All Passed feature is a printable report where the system shows / posts the examinee's performance after the examination. Lastly, the Print All Record feature is a printable report. The system generates all the examinees' performances: those who already took, those who did not, those who passed, those who failed, and those who need to retake the examination.

#### **4.4.2. Examinee Module**

For this software module, the Examinee login Interface feature allows the examinee to input his / her credentials, such as Examinee ID and Password. The Time Limit feature allows the examinee to be aware and to track the time allotted remaining for the whole examination period. The review Examination Items feature allows the examinee to correct/review/look back at their answers. In addition, the examinee will be given a limited time to review, provided there is time left to answer the missed questions before submission. Finally, the Confirmation submit button appears after reviewing the items. It is the message that prompts the examinee to correct/ change their answers before submission because once submitted, it cannot be undone and will be graded immediately.

#### **4.5 Discussion**

This research first identified the challenges of the institution, second, assessed the acceptability of the developed prototype, third, determined the level of accuracy of the developed system using regression algorithms, and lastly developed and designed features for the CPE. In comparison with related studies about machine learning models using different predictors like historical performances and other relevant factors that were used to predict student performances, this research achieved the integration of the development of computer-based examination and then predicts if the examinee who took the CPE is ready to take the actual examination. To generalize, this will help academic institutions and students to increase the probability of the examinees passing the licensure examination. As the results show, the institution is still using pen and paper modality while conducting the CPE. This research can help academic institutions with the same challenges and shift to a computer-based mode of conducting mock examinations. Regarding software engineering, several features can also be eliminated and added depending on the needs of the administrator/examinees. As for the predictive variables, predictors used in this study for the simple linear regression model and multilinear regression model can predict the actual licensure performance. This will serve as a basis for

crafting and considering other possible predictors and giving light to increase the accuracy level of the developed regression models.

## 5. CONCLUSION AND RECOMMENDATIONS

Based on the results and findings of the study, many ways or techniques can be used to predict student performance. In this case, regression models have been utilized in this study, and to build its model, some variables should be considered, like the predictors used in this study. Discussed below are the conclusions and recommendations based on the results of this study. For the challenges faced by the Institute of Engineering and Technology, the institution cannot monitor the examinees' weaknesses in the subjects or areas the students should have focused on for the incoming actual licensure examination. The institution should adapt to new technologies and use computer laboratories when conducting CPE to avoid more paperwork.

Using ISO 9126, the standard instrument to evaluate the software characteristics, the overall features of the proposed system were highly acceptable. It also means there is room for improvement, especially when using different operating systems and high-definition graphics where the user interface sometimes varies. This study used simple linear regression. It developed the equation with an accuracy level of 85 (round-off value), proving that the CPE system can predict the CAAP actual licensure examination results. On the other hand, the developed multilinear regression model and the equation are used with an accuracy level of 85 (round-off value) based on the predicted and actual data, proving that the 7 Technical Subject Areas results could predict the CAAP actual licensure examination results. It is recommended to use different algorithms and other machine learning techniques to compare how good is the model used in this study.

The developed CPE, together with the predictive models for the future or the actual licensure exam, is a helpful tool for academic institutions like PhilSCA to monitor and ensure that their examinees representing the said institution can get a high probability of passing on the actual CAAP AMT Licensure Examination. It is suggested to make the prototype more user-friendly and add options for the users to manage and explore the user interface for better user experience.

## 6. LIMITATIONS AND FUTURE STUDIES

A limitation of this study is the limited number of datasets used during data modeling, as this factor affects the level of accuracy for the machine learning algorithms. For the software prototype, the CPE system ran locally via a network server. It can only be accessed on-premise and cannot be accessed outside the institution's intranet to avoid leakages on results and sharing of answers to secure data integrity. In addition, future researchers interested in this type of study may also consider different predictors or factors that can contribute to modeling a good dataset for prediction. In the future, the software should have a Security mechanism and procedure to protect the CSV/XLSX files to ensure and avoid tampering with raw data.

### Acknowledgments

With profound indebtedness, the researcher would like to wholeheartedly express his earnest appreciation and gratefulness to the people who made this study possible, most especially to the following: Dr. Benilda Eleonor V. Comendador, his adviser, who helped him in crafting the idea of this study. His adviser's work also inspired him in academic performances on which he gained a deep interest in machine learning models; Captain Ronnie L. Briones, Department Manager of the Airmen Examination Board of Civil Aviation Authority of the Philippines and the Vice President for Academics of the Philippine State College of Aeronautics, Dr. Rodante G. Flores, for their generosity in allowing him to conduct this study; Dr Fabregas,

Assoc. Prof. Dela Fuente, Dr. Ado and Dr. Escobar, together with the panel of evaluators of this study, for their critique, comments, suggestions, and corrections to make this research better and factual. To His family for giving him a reason to dream and wake up positively each day. Above all, thanks and praises to Jehovah God for the gift of life and for giving him strength to carry on during his sleepless nights.

**Competing Interests:** The author declares that this work has no competing interests.

**Grant/Funding information:** The author declared that no grants supported this work

**Data Availability Statement:** The associated data is available upon reasonable request from the corresponding author.

**Declaration Statement of Generative AI:** The author of this work also used Python and Weka for data analysis and Machine learning development of the simple linear model and multilinear regression model. After using this tool, the author reviewed and verified the results of the prediction accuracy using the SAP Analytics Cloud tool as needed. The author takes (s) full responsibility for the publication's content.

## REFERENCES

- Abaya, S. A., Orig, D. A. D., & Montalbo, R. S. (2016). Using regression analysis in identifying students' performance in the board examination. *The Online Journal of New Horizons in Education*, 6(4), 290.
- Abass, O. A., Olajide, S. A., & Samuel, B. O. (2017). Development of web-based examination system using open source programming model. *Turkish Online Journal of Distance Education*, 18(2), 30-42. <https://doi.org/10.17718/tojde.306555>.
- Baars, G. J. A., Stijnen, T., & Splinter, T. A. W. (2017). A Model to Predict Student Failure in the First Year of the Undergraduate Medical Curriculum. *Health Professions Education*, 3(1), 5–14. <https://doi.org/10.1016/j.hpe.2017.01.001>.
- Broughton, S. (2017). *An evaluation of the effectiveness of a computer-aided assessment system for mathematics and engineering students* (Doctoral dissertation, Loughborough University).
- Chalmers, D., & McAusland, W. D. M. (2002). Computer-assisted assessment. *The handbook for economics lecturers*, 1-20. Retrieved from <http://www.economicsnetwork.ac.uk/handbook/caa/>
- Civil Aviation Authority of the Philippines. (2023). AC 02 Personnel Licensing. <https://Caap.Gov.Ph/Ac-02-Personnel-Licensing/>.
- Angeles, M. A. G. G. D. (2020). Predictors of Performance in Licensure Examination for Teachers. *Universal Journal of Educational Research*, 8(3), 835-843. DOI: 10.13189/ujer.2020.080314
- Samson, E. (2017). Computer-Based Examination System with Multifactor Authentication and Message Notification Features. *Faculty of Science Abubakar Tafawa Balewa University*. <https://doi.org/10.13140/RG.2.2.14713.88167>
- Hamsa, H., Indiradevi, S., & Kizhakkethottam, J. J. (2016). Student academic performance prediction model using decision tree and fuzzy genetic algorithm. *Procedia Technology*, 25, 326-332. <https://doi.org/10.1016/j.protcy.2016.08.114>
- Jain, R. (2016). Decision Support Systems: An Overview. *Research Gate*. [https://www.researchgate.net/publication/281085064\\_Decision\\_Support\\_Systems\\_an\\_Overview](https://www.researchgate.net/publication/281085064_Decision_Support_Systems_an_Overview)
- Jawaid, M., Moosa, F. A., Jaleel, F., & Ashraf, J. (2014). Computer-based assessment (CBA): Perception of residents at Dow University of Health Sciences. *Pakistan journal of medical sciences*, 30(4), 688. <https://doi.org/10.12669/pjms.304.5444>
- Kalechofsky, H. (2016). A simple framework for building predictive models. *A Little Data Science Business Guide*, 1-18. Retrieved from <http://www.msquared.com/wp-content/uploads/2017/01/A-Simple-Framework-for-Building-Predictive-Models.pdf>.

- Lu, O. H., Huang, A. Y., Huang, J. C., Lin, A. J., Ogata, H., & Yang, S. J. (2018). Applying learning analytics for the early prediction of Students' academic performance in blended learning. *Journal of Educational Technology & Society*, 21(2), 220-232. <https://www.jstor.org/stable/26388400>.
- Mesarić, J., & Šebalj, D. (2016). Decision trees for predicting the academic success of students. *Croatian Operational Research Review*, 7(2), 367-388. <https://doi.org/10.17535/crorr.2016.0025>
- Pandey, M., & Taruna, S. (2016). Towards the integration of multiple classifiers about the student's performance prediction. *Perspectives in science*, 8, 364-366. <https://doi.org/10.1016/j.pisc.2016.04.076>.
- Rakic, S., Tasic, N., Marjanovic, U., Softic, S., Lüftenegger, E., & Turcin, I. (2020). Student Performance on an E-Learning Platform: Mixed Method Approach. *International Journal of Emerging Technologies in Learning*, 15(2). <https://doi.org/10.3991/ijet.v15i02.11646>
- Rønning, F. (2017). Influence of computer-aided assessment on ways of working with mathematics. Teaching Mathematics and Its Applications: *International Journal of the IMA*, 36(2), 94-107. <https://doi.org/10.1093/teamat/hrx001>
- Shahiri, A. M., & Husain, W. (2015). A review on predicting student's performance using data mining techniques. *Procedia Computer Science*, 72, 414-422. <https://doi.org/10.1016/j.procs.2015.12.157>
- Sorour, S. E., Mine, T., Goda, K., & Hirokawa, S. (2015). A predictive model to evaluate student performance. *Journal of Information Processing*, 23(2), 192-201. <https://doi.org/10.2197/ipsjjip.23.192>
- Tarun, I. M. (2017). Prediction models for licensure examination performance using data mining classifiers for the online test and decision support system. *Asia Pacific Journal of Multidisciplinary Research*, 5(3), 10-21.
- Tarun, I. M., Gerardo, B. D., & Tanguilig III, B. T. (2014). Generating licensure examination performance models using PART and JRip classifiers: A data mining application in education. *International Journal of Computer and Communication Engineering*, 3(3), 203... <https://doi.org/10.7763/IJCCE.2014.V3.320>
- Umar, M. A., Ghazali, M., Saedudin, R. R., Ashraf, M., & Kasim, S. (2019). I am measuring software product quality using ISO 9126: A systematic review. In *International Journal of Advanced Trends in Computer Science and Engineering* (Vol. 8, Issue 1.3 S1, pp. 327–336). World Academy of Research in Science and Engineering. <https://doi.org/10.30534/ijatcse/2019/6081.32019>
- Zacharski, R. (2015). A Programmer's Guide to Data Mining: *The Ancient Art of the Numerati*.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations or the publisher, the editors and the reviewers. Any product that may be evaluated in this article or claim made by its manufacturer is not guaranteed or endorsed by the publisher.